Ethical, Legal, and Social Implications of Data Science Applications in Genomics Research

Kylie Hunter
Mentors: Shurjo Sen, Dave Kaufman

Purpose

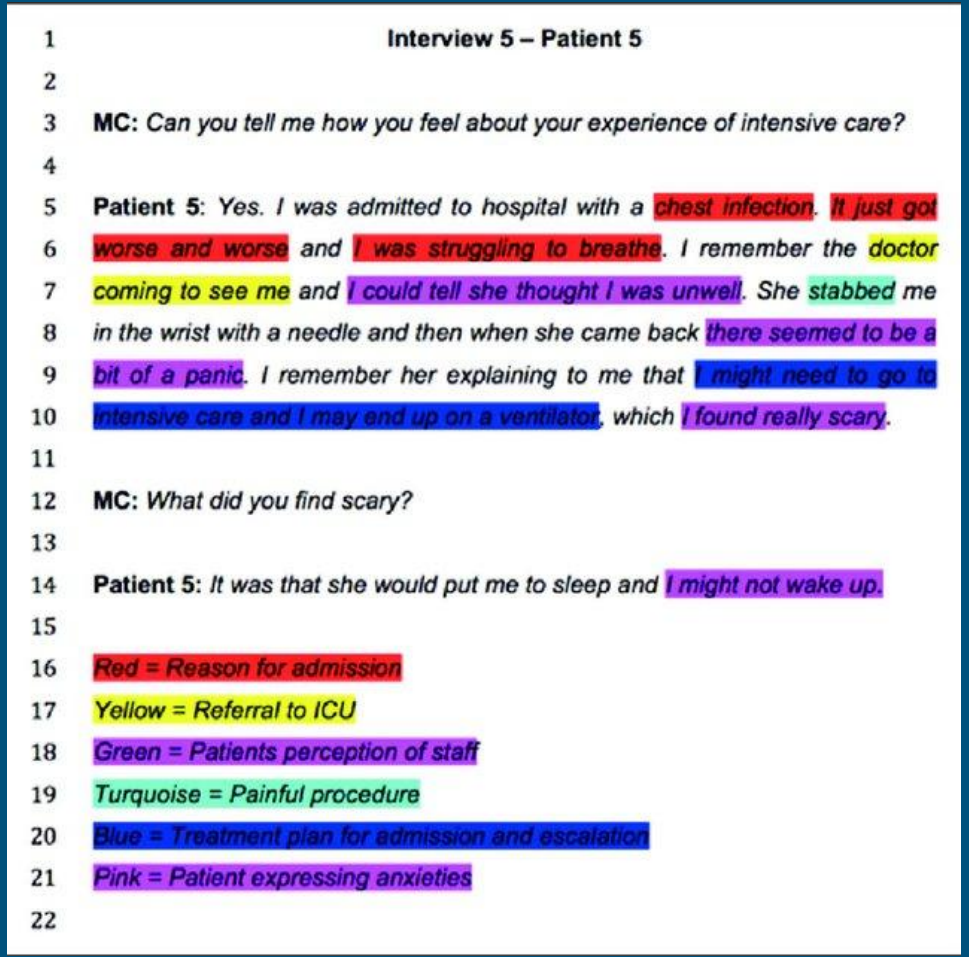
Identify ethical concerns surrounding the use of data science technologies in genomics

Highlight recommendations informed by experts to address these concerns

Methods

- Collect and analyze diverse expert opinions
- 1-hour structured interviews
- 12 interviewees, technical and nontechnical
  - NHGRI and outside professionals
- Transcript analysis

Example of analyzing (“coding”) a transcript:



Interview 5 – Patient 5

MC: *Can you tell me how you feel about your experience of intensive care?*

Patient 5: *Yes. I was admitted to hospital with a chest infection. It just got worse and worse and I was struggling to breathe. I remember the doctor coming to see me and I could tell she thought I was unwell. She stabbed me in the wrist with a needle and then when she came back there seemed to be a bit of a panic. I remember her explaining to me that I might need to go to intensive care and I may end up on a ventilator, which I found really scary.*

MC: *What did you find scary?*

Patient 5: *It was that she would put me to sleep and I might not wake up.*

Red = Reason for admission

Yellow = Referral to ICU

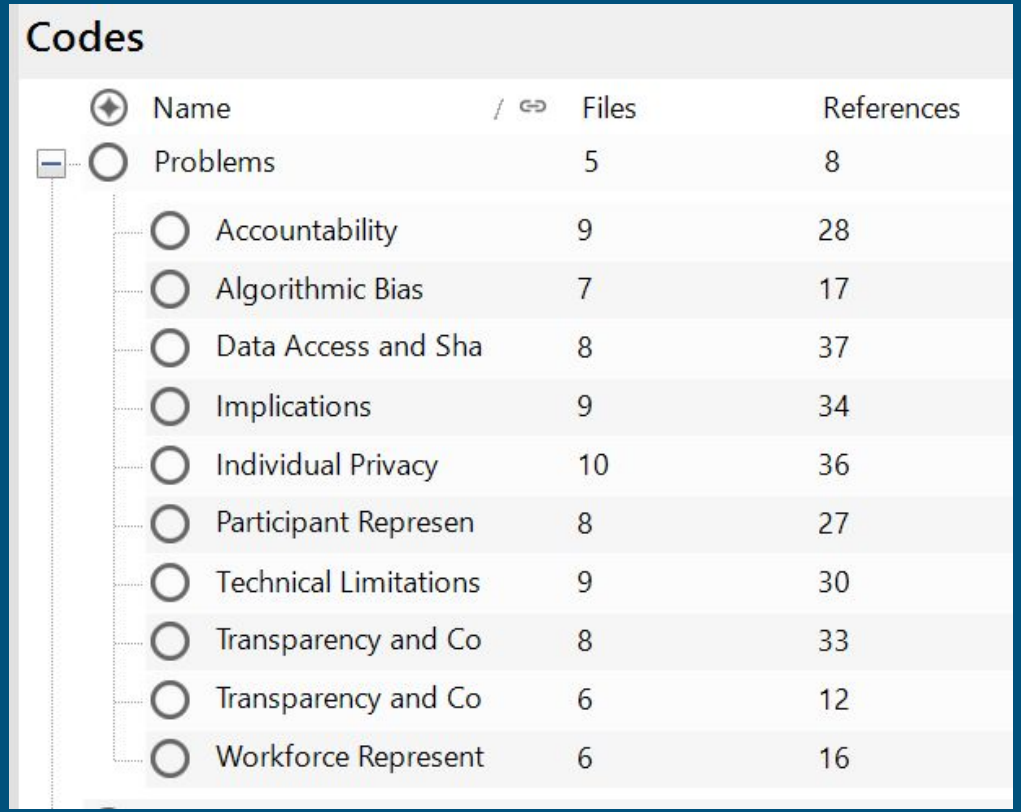
Green = Patients perception of staff

Turquoise = Painful procedure

Blue = Treatment plan for admission and escalation

Pink = Patient expressing anxieties

Example of Codes: ‘Problems’



Algorithmic Bias

Main Concerns

- Biased datasets
- Human bias
- Blind spots
  - Faulty analysis
  - “What” versus “why”
  - Overattribution of value
- Unknowns and Reproducibility
  - Validation/testing
  - Noise/overfitting

Transparency and Communication

Main Concerns

- Debates between privacy and open sharing
- Lack of collaboration
  - Social scientists
  - Community
- Data scientist involvement
  - Experiment design
  - Knowledge/accountability
- Presentation of results
  - Effect on communities

Some General Recommendations

- ★ Invest in building trust
  - Encourage participation and understanding
  - Reduce training bias
- ★ Culture of responsibility
  - Trainings with case studies
  - Consult with community
- ★ Transparency around data use/privacy
- ★ Outline expectations on ML use in research
  - Data auditing
- ★ List of questions
  - Ensure researchers are thinking about implications

Questions?